On the averaging principle for one-frequency systems. Seminorm estimates for the error

Carlo Morosi[1], Livio Pizzocchero[2]

[1] Dipartimento di Matematica, Politecnico di Milano,
P.za L. da Vinci 32, I-20133 Milano, Italy
e–mail: carmor@mate.polimi.it
[2] Dipartimento di Matematica, Università di Milano
Via C. Saldini 50, I-20133 Milano, Italy
and Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Italy
e–mail: livio.pizzocchero@mat.unimi.it

Abstract

We extend some previous results of our work [1] on the error of the averaging method, in the one-frequency case. The new error estimates apply to any separating family of seminorms on the space of the actions; they generalize our previous estimates in terms of the Euclidean norm. For example, one can use the new approach to get separate error estimates for each action coordinate. An application to rigid body under damping is presented. In a companion paper [2], the same method will be applied to the motion of a satellite around an oblate planet.

Keywords: Slow and fast motions, perturbations, averaging method.

AMS 2000 Subject classification: 70K65, 70K70, 34C29, 70H09, 37J40.

1 Introduction.

It is often stated by applied mathematicians that a good theorem on differential equations is one outlining a computational method for their solutions; if the method is approximate, quantitative error estimates should be provided.

In the case of ODEs with slow variables (”actions”) and fast angular variables, averaging over the angles is a well-known approximation technique; in the literature, the error of this method has been discussed mainly from a qualitative viewpoint, even in the simple case of one frequency (i.e., one angle only). The classical, qualitative estimates for this case (see e.g. [3]) have the form

$$\mathtt{I}(t) - \mathtt{J}(\varepsilon t) = O(\varepsilon) \qquad \text{for } t \in [0, O(1/\varepsilon)) \tag{1.1}$$

(uniformly in t) for a perturbation proportional to a parameter ε, in the limit $\varepsilon \to 0^+$; here, $\mathtt{I}(t)$ are the actions at time t, and $\mathtt{J}(\varepsilon t)$ is their approximation obtained from averaging (see paragraph 1A for more details).

In a previous paper [1], we have proposed in place of (1.1) a *fully quantitative* error estimate for the one-frequency averaging; this has the form

$$|\mathtt{I}(t) - \mathtt{J}(\varepsilon t)| \leqslant \varepsilon \mathfrak{n}(\varepsilon t) \qquad \text{for } t \in [0, U/\varepsilon) \ , \tag{1.2}$$

where $|\ |$ is the Euclidean norm on the space of the actions and $\mathfrak{n}$ is a computable function, determined by an integral inequality; U is a specified nonnegative constant, defining quantitatively the time interval where the estimate holds. (In fact, in some special cases considered in [1] the estimate holds even for very large values of U, e.g., $U \simeq 1/\varepsilon$). Let us repeat here a comment already done in the cited work: the idea of a really quantitative approach to the averaging methods has attracted little attention up to now, a notable exception being [4] that, in [1], we have briefly compared with our approach.

The present work is an improvement of [1] proposing more detailed error estimates, e.g., a separate bound on each component of the actions. These componentwise bounds are seen as a special case of a more general framework, where the estimates are expressed in terms of any separating family of seminorms on the space of the actions (a notion to be defined in the sequel). Our general estimates will take the form

$$|\mathtt{I}(t) - \mathtt{J}(\varepsilon t)|^{\mu} \leqslant \varepsilon \mathfrak{n}^{\mu}(\varepsilon t) \qquad \text{for } \mu \in M,\ t \in [0, U/\varepsilon) \ , \tag{1.3}$$

where $(|\ |^{\mu})_{\mu \in M}$ is the family of seminorms, labeled by a (finite) index set M.

To show the effectiveness of these bounds, in the present work we give a simple example related to rigid body dynamics. A more engaging application, concerning the motion of a satellite around an oblate planet, will be presented in the companion paper [2].

The forthcoming paragraphs 1A-1D introduce the following topics: the setting of [1] for one-frequency averaging, that we use partly in this paper; the new error estimates developed in the present work; the motivations to consider these refinements, and to formulate them in the language of seminorms; the organization of the paper.

1A. One-frequency averaging, in the framework of [1]. We consider an open set Λ of $\mathbf{R}^d$ and the one-dimensional torus $\mathbf{T}$ (referred to as the spaces of the actions and of the angular variable):

$$\Lambda = \{I = (I^i)_{i=1,\dots,d}\} \subset \mathbf{R}^d \ , \qquad \mathbf{T} := \mathbf{R}/2\pi\mathbf{Z} = \{\vartheta\} \ . \tag{1.4}$$

We suppose to be given a one-frequency system with a perturbation εf on the actions and εg on the angle: more precisely, we have a Cauchy problem

$$\begin{cases} d\mathtt{I}/dt = \varepsilon f(\mathtt{I}, \Theta) \ , & \mathtt{I}(0) = I_0 \ , \\ d\Theta/dt = \omega(\mathtt{I}) + \varepsilon g(\mathtt{I}, \Theta) \ , & \Theta(0) = \vartheta_0 \ , \end{cases} \tag{1.5}$$

under the assumptions

$$f = (f^i)_{i=1,\dots,d} \in C^m(\Lambda \times \mathbf{T}, \mathbf{R}^d), \ g \in C^m(\Lambda \times \mathbf{T}, \mathbf{R}), \ \omega \in C^m(\Lambda, \mathbf{R}) \ (m \geqslant 2) \ , \tag{1.6}$$

$$\omega(I) \neq 0 \text{ for all } I \in \Lambda \ , \quad I_0 \in \Lambda, \vartheta_0 \in \mathbf{T} \ , \qquad \varepsilon > 0 \ ;$$

the maximal solution (in the future) of (1.5) is a C^{m+1} function

$$(\mathtt{I}, \Theta) : [0, T) \to \Lambda \times \mathbf{T} \ , \qquad t \mapsto (\mathtt{I}(t), \Theta(t)) \ . \tag{1.7}$$

Throughout the paper, the initial data I_0, ϑ_0 and the perturbation parameter ε are fixed; for this reason, we do not indicate the dependence of $(\mathtt{I}, \Theta)$ and other functions on these objects. Needless to say, we are mainly interested in the case of small ε.

The averaged system associated to (1.5) is

$$\frac{d\mathtt{J}}{d\tau} = \overline{f}(\mathtt{J}) \ , \qquad \mathtt{J}(0) = I_0 \ , \tag{1.8}$$

$$\overline{f} = (\overline{f^i})_{i=1,\dots,d} \in C^m(\Lambda, \mathbf{R}^d) \ , \qquad I \mapsto \overline{f}(I) := \frac{1}{2\pi} \int_{\mathbf{T}} d\vartheta \ f(I, \vartheta) \ ;$$

the maximal solution (in the future) is a C^{m+1} function

$$\mathtt{J} : [0, W) \mapsto \Lambda \ , \qquad \tau \mapsto \mathtt{J}(\tau) \ . \tag{1.9}$$

The error of the averaging method is the function $t \mapsto \mathtt{I}(t) - \mathtt{J}(\varepsilon t)$ (defined whenever $\mathtt{I}(t)$ and $\mathtt{J}(\varepsilon t)$ exist); equivalently, one can consider the function

$$\mathtt{L} : t \mapsto \mathtt{L}(t) := \frac{1}{\varepsilon}[\mathtt{I}(t) - \mathtt{J}(\varepsilon t)] \ . \tag{1.10}$$

In [1] we have put the attention on the Euclidean norm

$$|\mathtt{L}(t)| = \sqrt{\sum_{i=1}^{d} \mathtt{L}^i(t)^2} \ ; \tag{1.11}$$

under natural conditions, we have derived for it a quantitative estimate

$$|\mathtt{L}(t)| \leqslant \mathfrak{n}(\varepsilon t) \qquad \text{for } t \in [0, U/\varepsilon) \tag{1.12}$$

(which is the same as (1.2)), where $\mathfrak{n} : [0, U) \to [0, +\infty)$ is a function determined by a fully explicit algorithm. To compute $\mathfrak{n}$, one must solve an integral inequality or a related differential equation on $[0, U)$, a task that in typical cases is performed numerically; however, for ε small this operation is much faster than the direct numerical solution of the perturbed system (1.5) for t in the long interval $[0, U/\varepsilon)$.

1B. Some variants in analyzing $\mathtt{L}(t)$. In view of applications, the following variants can be of interest:
(a) estimating a norm of $\mathtt{L}(t)$ different from (1.11);
(b) giving separate estimates on the absolute values $|\mathtt{L}^i(t)|$ of the components;
(c) considering a partition $\mathscr{P} = \{S, S'...\}$ of $\{1, ..., d\}$ into (nonempty) subsets $S, S', ...$ and estimating the components of $\mathtt{L}(t)$ in each subset: for example, one could analyze the quantities

$$\sqrt{\sum_{i \in S} \mathtt{L}^i(t)^2} \ , \ \sqrt{\sum_{i \in S'} \mathtt{L}^i(t)^2} \ , ... \ . \tag{1.13}$$

Here are some reasons to study each component separately, or to group them into subsets: the components could measure physically nonhomogeneous quantities; one expects relevant differences in their numerical values, even in the orders of magnitude. All these facts will occur in the example of Section 3, related to rigid body dynamics.

1C. General estimates for $\mathtt{L}(t)$ via seminorms. A unified way to treat (a) (b) (c) and other situations is to consider on $\mathbf{R}^d$ a *separating family of seminorms*, and use them to estimate $\mathtt{L}(t)$. Let us recall that a seminorm on $\mathbf{R}^d$ is a map

$$\mathbf{R}^d \to [0, +\infty) \ , \qquad X \mapsto |X| \ , \tag{1.14}$$

homogeneous and subadditive:

$$|\lambda X| = |\lambda||X| \ , \qquad |X + Y| \leqslant |X| + |Y| \qquad \text{for } X, Y \in \mathbf{R}^d,\ \lambda \in \mathbf{R} \tag{1.15}$$

($|\lambda|$ is the absolute value of λ; the first relation, with $\lambda = 0$, gives $|0| = 0$). An example of a seminorm is the function $|\ |^i$ on $\mathbf{R}^d$, where i is any integer in $\{1, ..., d\}$ and

$$|X|^i := |X^i| \tag{1.16}$$

for all $X = (X^1, ..., X^d) \in \mathbf{R}^d$; more generally, if S is a (nonempty) subset of $\mathbf{R}^d$ we can define a seminorm $| \, |^S$ on $\mathbf{R}^d$, setting

$$|X|^S := \sqrt{\sum_{i \in S} (X^i)^2} \ . \tag{1.17}$$

A *norm* on $\mathbf{R}^d$ can be defined as a seminorm with the supplementary *separation* property

$$|X| = 0 \quad \Rightarrow \quad X = 0 \ . \tag{1.18}$$

Clearly, this property is lacking (for $d > 1$) in the example (1.16); it is also lacking in (1.17), unless $S = \{1, ..., d\}$. However, these examples with variable i or S, and other situations, carry to *families* of seminorms possessing the separation property in a collective sense. To be precise, a *separating family of seminorms* on $\mathbf{R}^d$ is a family $(| \, |^\mu)_{\mu \in M}$, where M is a finite set, such that $| \, |^\mu$ is a seminorm for each μ and, for all $X \in \mathbf{R}^d$,

$$|X|^\mu = 0 \quad \text{for each } \mu \in M \quad \Rightarrow \quad X = 0 \ . \tag{1.19}$$

An example of a separating family is formed by all the seminorms (1.16), with i ranging in $\{1, ..., d\}$. Another example is the family (1.17), labeled by the subsets S in a partition $\mathscr{P}$ of $\{1, ..., d\}$.

Throughout the paper, our estimates for $\mathtt{L}(t)$ will concern the nonnegative quantities

$$|\mathtt{L}(t)|^\mu \qquad (\mu \in M) \tag{1.20}$$

for any chosen separating family of seminorms on $\mathbf{R}^d$. Case (a) of the previous paragraph corresponds to the choice $M = \{1\}$ and $| \, |^1 =$ a norm $| \, |$ on $\mathbf{R}^d$; case (b) corresponds to the family (1.16) with $M = \{1, ..., d\}$, and case (c) to the family (1.17) with $M = \mathscr{P}$.

1D. Organization of the paper. Section 2 is the main body of the paper: after recalling a basic Lemma from [1], we construct the general framework to estimate $\mathtt{L}$ through a separating family of seminorms. The conclusion is a set of inequalities

$$|\mathtt{L}(t)|^\mu \leqslant \mathfrak{n}^\mu(\varepsilon t) \qquad \text{for } \mu \in M, \, t \in [0, U/\varepsilon) \tag{1.21}$$

(i.e., of the form (1.3)), where the estimators $\mathfrak{n}^\mu : [0, U) \to [0, +\infty)$ are determined solving a system of integral inequalities (Proposition 2.6), or of differential equations related to them (Proposition 2.7). Section 3 presents an example, arising from the dynamics of a rigid body under damping; this was introduced in [1] and will be reconsidered from the present viewpoint, deriving separate error estimates for each one of the two actions. The Appendices A, B contain the proofs of the previously mentioned Propositions.

In spite of the frequent reference to [1], in writing the present paper we have tried to make it reasonably self-contained.

2 Main results.

2A. Notations. (i) Throughout the paper, vectors of $\mathbf{R}^d$ are written with upper indices: $X = (X^i)_{i=1,...,d}$, as already done in the Introduction. Due to the fact that $\mathbf{R}^d$ has a canonical basis, the tensors on $\mathbf{R}^d$ of any type (p, q) can be identified with tables of real numbers, that we write in the usual style with p upper and q lower indices. In the sequel we will often use the tensor spaces

$$\begin{aligned} \mathrm{T}^1_1(\mathbf{R}^d) &= \{\mathscr{A} = (\mathscr{A}^i_j) \mid \mathscr{A}^i_j \in \mathbf{R} \text{ for } i, j = 1, ..., d\} \ , \\ \mathrm{T}^2_0(\mathbf{R}^d) &= \{\mathscr{B} = (\mathscr{B}^{ij}) \mid \mathscr{B}^{ij} \in \mathbf{R} \text{ for } i, j = 1, ..., d\} \ , \\ \mathrm{T}^1_2(\mathbf{R}^d) &= \{\mathscr{C} = (\mathscr{C}^i_{jk}) \mid \mathscr{C}^i_{jk} \in \mathbf{R} \text{ for } i, j, k = 1, ..., d\} \ . \end{aligned} \tag{2.1}$$

We use systematically Einstein's summation convention on repeated upper and lower indices. Let $X, Y \in \mathbf{R}^d$, $\mathscr{A}, \mathscr{D} \in \mathrm{T}^1_1(\mathbf{R}^d)$ and $\mathscr{C} \in \mathrm{T}^1_2(\mathbf{R}^d)$; then, $\mathscr{A}X \in \mathbf{R}^d$ and $\mathscr{C}XY \in \mathbf{R}^d$ are the vectors of components $(\mathscr{A}X)^i = \mathscr{A}^i_j X^j$, $(\mathscr{C}XY)^i = \mathscr{C}^i_{jk} X^j Y^k$; $\mathscr{A}\mathscr{D}, \mathscr{C}X \in \mathrm{T}^1_1(\mathbf{R}^d)$ are the tensors of components $(\mathscr{A}\mathscr{D})^i_k = \mathscr{A}^i_j \mathscr{D}^j_k$, $(\mathscr{C}X)^i_k = \mathscr{C}^i_{jk} X^j$.

(ii) We fix on $\mathbf{R}^d$ a separating family of seminorms

$$| \ |^\mu \qquad (\mu \in M) \ , \tag{2.2}$$

with M a finite set. To go on, we need some seminorm families on the tensor spaces $\mathrm{T}^1_1(\mathbf{R}^d)$ and $\mathrm{T}^1_2(\mathbf{R}^d)$; of course a seminorm on $\mathrm{T}^1_1(\mathbf{R}^d)$ is a homogeneous, subadditive map

$$| \ | : \mathrm{T}^1_1(\mathbf{R}^d) \to [0, +\infty) \ , \qquad \mathscr{A} \mapsto |\mathscr{A}| \tag{2.3}$$

and a seminorm on $\mathrm{T}^1_2(\mathbf{R}^d)$ is defined similarly.

Keeping fixed the family (2.2), a *consistent family of seminorms* on $\mathrm{T}^1_1(\mathbf{R}^d)$ is one of the form

$$| \ |^\mu_\nu \qquad (\mu, \nu \in M) \ , \tag{2.4}$$

with the property

$$|\mathscr{A}X|^\mu \leqslant |\mathscr{A}|^\mu_\nu \, |X|^\nu \qquad \text{for all } \mathscr{A} \in \mathrm{T}^1_1(\mathbf{R}^d), \ X \in \mathbf{R}^d \text{ and } \mu \in M \tag{2.5}$$

(here and in the sequel, Einstein's summation convention is also employed for repeated indices with values in M). Similarly, a *consistent family of seminorms* on $\mathrm{T}^1_2(\mathbf{R}^d)$ is a family of seminorms

$$| \ |^\mu_{\nu\kappa} \qquad (\mu, \nu, \kappa \in M) \ , \tag{2.6}$$

such that

$$|\mathscr{C}XY|^\mu \leqslant |\mathscr{C}|^\mu_{\nu\kappa} |X|^\nu |Y|^\kappa \qquad \text{for all } \mathscr{C} \in \mathrm{T}^1_1(\mathbf{R}^d), \ X, Y \in \mathbf{R}^d \text{ and } \mu \in M \ . \tag{2.7}$$

The existence of such consistent families can be proved using the separation property of (2.2) ([1]). Of course, if we use on $\mathbf{R}^d$ the seminorms $|\ |^i$ of Eq. (1.16) we have on $\mathrm{T}^1_1(\mathbf{R}^d)$ and $\mathrm{T}^1_2(\mathbf{R}^d)$ the following consistent families of seminorms, also taking the absolute values of the tensor components:

$$|\mathscr{A}|^i_j := |\mathscr{A}^i_j| \ , \qquad |\mathscr{C}|^i_{jk} := |\mathscr{C}^i_{jk}| \qquad (i,j,k = 1,...,d) \ . \tag{2.8}$$

(iii) In the sequel we intend

$$\Lambda_\dagger := \{(I, \delta I) \in \Lambda \times \mathbf{R}^d \mid [I, I + \delta I] \subset \Lambda\} \ , \tag{2.9}$$

where $[I, I+\delta I]$ is the closed segment in $\mathbf{R}^d$ with the indicated extremes.

2B. The integral equation for $\mathtt{L}$. We consider the perturbed and averaged systems (1.5) (1.8), for fixed $\varepsilon > 0$ and initial data I_0, ϑ_0. We introduce the functions $s \in C^m(\Lambda \times \mathbf{T}, \mathbf{R}^d)$ and $p \in C^{m-1}(\Lambda \times \mathbf{T}, \mathbf{R}^d)$ such that

$$f = \overline{f} + \omega \, \frac{\partial s}{\partial \vartheta} \ , \quad \overline{s} = 0 \ ; \qquad p := \frac{\partial s}{\partial I} f + \frac{\partial s}{\partial \vartheta} g \ ; \tag{2.10}$$

these equations, and the forthcoming ones are always understood in the tensorial sense ([2]).

From now on, U stands for an element of $(0, +\infty]$.

2.1 *Lemma.* *Suppose the solution $\mathtt{J}$ of (1.8) exists for $\tau \in [0, U)$. Denote by $\mathtt{R} : [0, U) \to T^1_1(\boldsymbol{R}^d)$, $\tau \mapsto \mathtt{R}(\tau)$ and $\mathtt{K} : [0, U) \to \boldsymbol{R}^d$, $\tau \mapsto \mathtt{K}(\tau)$ the solutions of*

$$\frac{d\mathtt{R}}{d\tau} = \frac{\partial \overline{f}}{\partial I}(\mathtt{J})\, \mathtt{R} \ , \qquad \mathtt{R}(0) = 1_d \ ; \tag{2.11}$$

$$\frac{d\mathtt{K}}{d\tau} = \frac{\partial \overline{f}}{\partial I}(\mathtt{J})\, \mathtt{K} + \overline{p}(\mathtt{J}) \ , \qquad \mathtt{K}(0) = 0 \tag{2.12}$$

---

[1] This follows from much more general results on multilinear maps and seminorms that can be found, e.g., in [5].

[2] For better clarity, let us give only some examples. The equivalents in components of Eq. (2.10) are

$$f^i = \overline{f^i} + \omega \, \frac{\partial s^i}{\partial \vartheta} \ , \quad \overline{s^i} = 0 \ ; \qquad p^i := \frac{\partial s^i}{\partial I^j} f^j + \frac{\partial s^i}{\partial \vartheta} g \ .$$

In the forthcoming Eq.s (2.18) and (2.20), the relations about for $\mathscr{M}, \overline{f}$ and $\overline{f}$, $\mathscr{H}$ mean, respectively:

$$\mathscr{M}^i_k := \frac{\partial^2 \overline{f^i}}{\partial I^j \partial I^k} \, \overline{f^j} - \frac{\partial \overline{f^i}}{\partial I^j} \, \frac{\partial \overline{f^j}}{\partial I^k} \ ;$$

$$\overline{f}^i(I + \delta I) = \overline{f}^i(I) + \frac{\partial \overline{f}^i}{\partial I^j}(I) \delta I^j + \frac{1}{2} \mathscr{H}^i_{jk}(I, \delta I) \delta I^j \delta I^k \ .$$

(these exist and are C^m; $\mathtt{R}(\tau)$ is an invertible matrix for all $\tau \in [0, U)$, and $\mathtt{K}(\tau) = \mathtt{R}(\tau) \int_0^\tau d\tau' \, \mathtt{R}(\tau')^{-1} \overline{p}(\mathtt{J}(\tau'))$. For $d = 1$, $\mathtt{R}(\tau) = \exp \int_0^\tau d\tau' \, \dfrac{\partial \overline{f}}{\partial I}(\mathtt{J}(\tau')) \in (0, +\infty)$).

Furthermore, assume that the solution $(\mathtt{I}, \Theta)$ of the perturbed system (1.5) exists for $t \in [0, U/\varepsilon)$, with $(\mathtt{J}(\varepsilon t), \mathtt{I}(t) - \mathtt{J}(\varepsilon t)) \in \Lambda_\dagger$. Finally, define

$$\mathtt{L} : [0, U/\varepsilon) \to \boldsymbol{R}^d \ , \qquad t \mapsto \mathtt{L}(t) := \frac{1}{\varepsilon} \left[\mathtt{I}(t) - \mathtt{J}(\varepsilon t) \right] \ . \tag{2.13}$$

Then, for $t \in [0, U/\varepsilon)$,

$$\begin{aligned} \mathtt{L}(t) = s(\mathtt{I}(t), \Theta(t)) - \mathtt{R}(\varepsilon t) \, s(I_0, \vartheta_0) - \mathtt{K}(\varepsilon t) \\ -\varepsilon \Big(w(\mathtt{I}(t), \Theta(t)) - \frac{\partial \overline{f}}{\partial I}(\mathtt{J}(\varepsilon t)) \, v(\mathtt{I}(t), \Theta(t)) \Big) \\ + \varepsilon^2 \mathtt{R}(\varepsilon t) \int_0^t dt' \, \mathtt{R}^{-1}(\varepsilon t') \Big(u(\mathtt{I}(t'), \Theta(t')) - \frac{\partial \overline{f}}{\partial I}(\mathtt{J}(\varepsilon t'))(w + q)(\mathtt{I}(t'), \Theta(t')) \\ - \mathscr{M}(\mathtt{J}(\varepsilon t')) v(\mathtt{I}(t'), \Theta(t')) - \mathscr{G}(\mathtt{J}(\varepsilon t'), \varepsilon \mathtt{L}(t')) \mathtt{L}(t') + \frac{1}{2} \mathscr{H}(\mathtt{J}(\varepsilon t'), \varepsilon \mathtt{L}(t')) \, \mathtt{L}(t')^2 \Big) \ . \end{aligned} \tag{2.14}$$

In the above, $v \in C^m(\Lambda \times \boldsymbol{T}, \boldsymbol{R}^d)$, $q, w \in C^{m-1}(\Lambda \times \boldsymbol{T}, \boldsymbol{R}^d)$, $u \in C^{m-2}(\Lambda \times \boldsymbol{T}, \boldsymbol{R}^d)$ and $\mathscr{M} \in C^{m-2}(\Lambda, T^1_1(\boldsymbol{R}^d))$ are the functions uniquely defined by the following equations:

$$s = \omega \frac{\partial v}{\partial \vartheta} \ , \qquad v(I, \vartheta_0) = 0 \quad \textit{for all } I \in \Lambda; \tag{2.15}$$

$$q := \frac{\partial v}{\partial I} f + \frac{\partial v}{\partial \vartheta} g \ ; \tag{2.16}$$

$$p = \overline{p} + \omega \frac{\partial w}{\partial \vartheta} \ , \qquad w(I, \vartheta_0) = 0 \quad \textit{for all } I \in \Lambda \ ; \tag{2.17}$$

$$u := \frac{\partial w}{\partial I} f + \frac{\partial w}{\partial \vartheta} g \ ; \qquad \mathscr{M} := \frac{\partial^2 \overline{f}}{\partial I^2} \, \overline{f} - \left(\frac{\partial \overline{f}}{\partial I} \right)^2 \ . \tag{2.18}$$

Furthermore, $\mathscr{G} \in C^{m-2}(\Lambda_\dagger, T^1_1(\boldsymbol{R}^d))$ and $\mathscr{H} \in C^{m-2}(\Lambda_\dagger, T^1_2(\boldsymbol{R}^d))$ are two functions such that, for all $(I, \delta I) \in \Lambda_\dagger$,

$$\overline{p}(I + \delta I) = \overline{p}(I) + \mathscr{G}(I, \delta I) \delta I \ , \tag{2.19}$$

$$\overline{f}(I + \delta I) = \overline{f}(I) + \frac{\partial \overline{f}}{\partial I}(I) \delta I + \frac{1}{2} \mathscr{H}(I, \delta I) \delta I^2 \ , \quad \mathscr{H}^i_{jk}(I, \delta I) = \mathscr{H}^i_{kj}(I, \delta I) \ . \tag{2.20}$$

Proof. See [1]. □

2.2 Remark. In dimension $d = 1$, Eqs. (2.19) (2.20) can be uniquely solved for $\mathscr{G}(I, \delta I)$ and $\mathscr{H}(I, \delta I)$; in any dimension we have the solutions given by Taylor's formula, i.e.,

$$\mathscr{G}(I, \delta I) := \int_0^1 dx\, \frac{\partial \overline{p}}{\partial I}(I+x\delta I)\ , \quad \mathscr{H}(I, \delta I) := 2 \int_0^1 dx\, (1-x) \frac{\partial^2 \overline{f}}{\partial I^2}(I+x\delta I)\ . \quad (2.21)$$

If $\overline{p}$ (resp. $\overline{f}$) is a polynomial or rational function of the actions, $\mathscr{G}$ (resp. $\mathscr{H}$) can be obtained in a simpler way by direct inspection of Eq. (2.19) (resp. (2.20)).

Now, from the integral equation (2.14) for the function $t \mapsto \mathtt{L}(t)$ we wish to infer a system of integral inequalities for the functions $t \mapsto |\mathtt{L}(t)|^\mu$, where $(|\ |^\mu)$ is any separating family of seminorms on $\mathbf{R}^d$. This requires a set of auxiliary functions, estimating several characters in (2.14), which are introduced hereafter.

2C. New auxiliary functions. For each set Z, we write

$$Z^M := \{z = (z^\mu)_{\mu \in M} \mid z^\mu \in Z \ \ \forall\, \mu\}\ . \quad (2.22)$$

For $J \in \mathbf{R}^d$ and $\varrho = (\varrho^\mu) \in [0, +\infty]^M$, we put

$$B(J, \varrho) := \{I \in \mathbf{R}^d \mid |I - J|^\mu < \varrho^\mu \ \ \forall\, \mu \in M\}\ . \quad (2.23)$$

We further assume the following.
(i) $\rho = (\rho^\mu) \in C([0, U), [0, +\infty]^M)$ is a function such that

$$B(\mathtt{J}(\tau), \rho(\tau)) \subset \Lambda \qquad \text{for } \tau \in [0, U)\ . \quad (2.24)$$

We put

$$\Gamma_\rho := \{(\tau, r) \ \in [0, U) \times [0, +\infty)^M \mid r^\mu < \rho^\mu(\tau)\ \forall\, \mu \in M\}\ . \quad (2.25)$$

(ii) $a^\mu, b^\mu, c^\mu, d^\mu_\nu, e^\mu_{\nu\kappa} \in C(\Gamma_\rho, [0, +\infty))$ $(\mu, \nu, \kappa \in M)$ are functions such that for any $\tau \in [0, U)$, $\delta J \in B(0, \rho(\tau))$ and $\vartheta \in \mathbf{T}$,

$$|s(\mathtt{J}(\tau) + \delta J, \vartheta) - \mathtt{R}(\tau) s(I_0, \vartheta_0) - \mathtt{K}(\tau)|^\mu \leqslant a^\mu(\tau, |\delta J|)\ , \quad (2.26)$$

$$\Big| w(\mathtt{J}(\tau) + \delta J, \vartheta) - \frac{\partial \overline{f}}{\partial I}(\mathtt{J}(\tau))\, v(\mathtt{J}(\tau) + \delta J, \vartheta) \Big|^\mu \leqslant b^\mu(\tau, |\delta J|)\ , \quad (2.27)$$

$$\Big| u(\mathtt{J}(\tau) + \delta J, \vartheta) - \frac{\partial \overline{f}}{\partial I}(\mathtt{J}(\tau))(w + q)(\mathtt{J}(\tau) + \delta J, \vartheta) \quad (2.28)$$

$$- \mathscr{M}(\mathtt{J}(\tau)) v(\mathtt{J}(\tau) + \delta J, \vartheta) \Big|^\mu \leqslant c^\mu(\tau, |\delta J|)\ ,$$

$$|\mathscr{G}(\mathtt{J}(\tau), \delta J)|^\mu_\nu \leqslant d^\mu_\nu(\tau, |\delta J|)\ , \quad (2.29)$$

$$|\mathscr{H}(\mathtt{J}(\tau), \delta J)|^\mu_{\nu\kappa} \leqslant e^\mu_{\nu\kappa}(\tau, |\delta J|)\ . \quad (2.30)$$

In the above, one always intends

$$|\delta J| := (|\delta J|^{\lambda})_{\lambda \in M} \ . \tag{2.31}$$

The functions $c^{\mu}, d^{\mu}_{\nu}, e^{\mu}_{\nu\kappa}$ are assumed to be nondecreasing with respect to the variable r, i.e.,

$$(\tau, r), (\tau, r') \in \Gamma_{\rho}, \quad r^{\lambda} \leqslant r'^{\lambda} \ \forall \lambda \in M \quad \Rightarrow \quad c^{\mu}(\tau, r) \leqslant c^{\mu}(\tau, r') \tag{2.32}$$

and similarly for d^{μ}_{ν} and $e^{\mu}_{\nu\kappa}$. Given $a^{\mu}, ..., e^{\mu}_{\nu\kappa}$, we define the functions

$$\alpha^{\mu} \in C(\Gamma_{\rho}, [0, +\infty)), \quad \alpha^{\mu}(\tau, r) := a^{\mu}(\tau, r) + \varepsilon b^{\mu}(\tau, r) \ , \tag{2.33}$$

$$\gamma^{\mu} \in C(\Gamma_{\rho} \times [0, +\infty)^{M}, [0, +\infty)), \tag{2.34}$$

$$\gamma^{\mu}(\tau, r, \ell) := c^{\mu}(\tau, r) + d^{\mu}_{\nu}(\tau, r)\ell^{\nu} + \frac{1}{2} e^{\mu}_{\nu\kappa}(\tau, r)\ell^{\nu}\ell^{\kappa} \ .$$

In the sequel we will set $\alpha := (\alpha^{\mu}) \in C(\Gamma_{\rho}, [0, +\infty)^{M})$, and intend γ similarly.
(iii) R^{μ}_{ν}, $P^{\mu}_{\nu} \in C([0, U), [0, +\infty))$ $(\mu, \nu \in M)$ are functions such that, for $\tau \in [0, U)$,

$$|\mathtt{R}(\tau)|^{\mu}_{\nu} \leqslant R^{\mu}_{\nu}(\tau) \ , \quad |\mathtt{R}^{-1}(\tau)|^{\mu}_{\nu} \leqslant P^{\mu}_{\nu}(\tau) \ . \tag{2.35}$$

2.3 Remarks. **(a)** In the main following statements about $|\mathtt{L}(t)|^{\mu}$ (Propositions 2.6, 2.7), the functions $b^{\mu}, ..., e^{\mu}_{\nu\kappa}$ will always be multiplied by the small factor ε. For this reason, in applications one can determine $b^{\mu}, ..., e^{\mu}_{\nu\kappa}$ via fairly rough majorizations of the left-hand sides of Eqs. (2.27)-(2.30). The situation is different for the functions a^{μ}, that are not multiplied by ε and so require accurate estimates.
(b) A trivial choice for the functions in (iii) is $R^{\mu}_{\nu}(\tau) := |\mathtt{R}(\tau)|^{\mu}_{\nu}$, $P^{\mu}_{\nu}(\tau) := |\mathtt{R}^{-1}(\tau)|^{\mu}_{\nu}$. This is not satisfactory if one wants more than the C^{0} regularity: in fact, this choice does not grant R^{μ}_{ν} and P^{μ}_{ν} to be C^{k} for any $k \geqslant 1$. On the other hand, C^{k} regularity with $k = 1$ or 2 will be required by some subsequent manipulations, and in view of this we leave R^{μ}_{ν} and P^{μ}_{ν} unspecified.

2D. Integral inequalities for $(|\mathtt{L}|^{\mu})$**.** We keep the assumptions and notations of the previous paragraph.

2.4 *Lemma.* *Assume that the solution* $(\mathtt{I}, \Theta)$ *of the perturbed system exists on* $[0, U/\varepsilon)$ *and that* $|\mathtt{L}(t)|^{\mu} < \rho^{\mu}(\varepsilon t)/\varepsilon$ *for all* $\mu \in M$, $t \in [0, U/\varepsilon)$*, Then, for all* μ *and* t *as above,*

$$|\mathtt{L}(t)|^{\mu} \leqslant \alpha^{\mu}(\varepsilon t, \varepsilon|\mathtt{L}(t)|) + \varepsilon^{2} R^{\mu}_{\lambda}(\varepsilon t) \int_{0}^{t} dt' P^{\lambda}_{\kappa}(\varepsilon t') \, \gamma^{\kappa}(\varepsilon t', \varepsilon|\mathtt{L}(t')|, |\mathtt{L}(t')|) \ , \tag{2.36}$$

intending $|\mathtt{L}(t)| := (|\mathtt{L}(t)|^{\nu})_{\nu \in M}$.

Proof. We take the μ-th seminorm of both sides in Eq. (2.14). To estimate the right-hand side, we use the consistency inequalities (2.5) (2.7), together with the relation $| \int_{0}^{t} dt' .. |^{\lambda} \leqslant \int_{0}^{t} dt' |..|^{\lambda}$; next, we apply the inequalities (2.26)–(2.30) with $\delta J = \mathtt{I}(t) - \mathtt{J}(\varepsilon t) = \varepsilon \mathtt{L}(t)$, and the inequalities (2.35). In this way we obtain

$$|\mathtt{L}(t)|^\mu \leqslant a^\mu(\varepsilon t, \varepsilon|\mathtt{L}(t)|\,) + \varepsilon\, b^\mu(\varepsilon t, \varepsilon|\mathtt{L}(t)|\,) + \varepsilon^2 R^\mu_\lambda(\varepsilon t) \int_0^t dt' P^\lambda_\kappa(\varepsilon t') \tag{2.37}$$

$$\times \Big(c^\kappa(\varepsilon t', \varepsilon|\mathtt{L}(t')|) + d^\kappa_\nu(\varepsilon t', \varepsilon|\mathtt{L}(t')|)\, |\mathtt{L}(t')|^\nu + \frac{1}{2}\, e^\kappa_{\nu\varsigma}(\varepsilon t', \varepsilon|\mathtt{L}(t')|)\, |\mathtt{L}(t')|^\nu |\mathtt{L}(t')|^\varsigma \Big) \ .$$

Now, the thesis (2.36) follows from the definitions (2.33),(2.34) of α, γ. $\square$

2E. A general fact on integral inequalities. This result is stated without proof, being a simple variation of similar ones appearing in [1] [6].

2.5 *Lemma.* *Let $T \in (0, +\infty]$, $\delta = (\delta^\mu) \in C([0,T), [0,+\infty]^M)$ and*

$$\Xi := \{(t,\ell) \in [0,T) \times [0,+\infty)^M \mid \ell^\mu < \delta^\mu(t)\ \forall\, \mu\ \}, \tag{2.38}$$

$$\mathrm{H} := \{(t,t',\ell) \mid t \in [0,T),\ t' \in [0,t],\ (t',\ell) \in \Xi\ \} \ .$$

Consider two functions $\xi = (\xi^\mu) \in C(\Xi, [0,+\infty)^M)$ and $\eta = (\eta^\mu) \in C(\mathrm{H}, [0,+\infty)^M)$. Let each function η^μ be nondecreasing in the last variable: $\eta^\mu(t,t',\ell') \leqslant \eta^\mu(t,t',\ell)$ if $(t,t',\ell), (t,t',\ell') \in \mathrm{H}$ and $\ell'^\nu \leqslant \ell^\nu$ for all $\nu \in M$; furthermore, let $\mathfrak{l} = (\mathfrak{l}^\mu), \mathfrak{v} = (\mathfrak{v}^\mu) \in C([0,T), [0,+\infty)^M)$ be such that graph $\mathfrak{l}$, graph $\mathfrak{v} \subset \Xi$, and

$$\mathfrak{l}^\mu(0) = 0 \ , \qquad \mathfrak{l}^\mu(t) \leqslant \xi^\mu(t, \mathfrak{l}(t)) + \int_0^t dt' \eta^\mu(t,t',\mathfrak{l}(t')) \ , \tag{2.39}$$

$$\mathfrak{v}^\mu(t) > \xi^\mu(t, \mathfrak{v}(t)) + \int_0^t dt' \eta^\mu(t,t',\mathfrak{v}(t')) \tag{2.40}$$

for all $\mu \in M$, $t \in [0,T)$. Then, for all such μ and t,

$$\mathfrak{l}^\mu(t) < \mathfrak{v}^\mu(t) \ . \tag{2.41}$$

2F. The main Proposition. We still assume that the solution $\mathtt{J}$ of the averaged system exists on $[0,U)$, and define $\mathtt{R}, \mathtt{K}$ via Eqs. (2.11) (2.12). Moreover, let us be given a set of functions $\rho^\mu, a^\mu, b^\mu, c^\mu, d^\mu_\nu, e^\mu_{\nu\kappa}$ as in paragraph 2C; α^μ and γ^μ are defined consequently, as indicated therein.

2.6 *Proposition.* *Assume there is a function $\mathfrak{n} = (\mathfrak{n}^\mu) \in C([0,U), [0,+\infty)^M)$ such that, for all $\mu \in M$ and $\tau \in [0,U)$,*

$$\mathfrak{n}^\mu(\tau) < \rho^\mu(\tau)/\varepsilon \ , \tag{2.42}$$

$$\mathfrak{n}^\mu(\tau) > \alpha^\mu(\tau, \varepsilon\mathfrak{n}(\tau)) + \varepsilon R^\mu_\lambda(\tau) \int_0^\tau d\tau' P^\lambda_\nu(\tau')\, \gamma^\nu(\tau', \varepsilon\mathfrak{n}(\tau'), \mathfrak{n}(\tau')) \ . \tag{2.43}$$

Then, the solution $(\mathtt{I}, \Theta)$ of the perturbed system exists on $[0, U/\varepsilon)$; furthermore, defining $\mathtt{L}$ as in Eq. (2.13) we have

$$|\mathtt{L}(t)|^\mu < \mathfrak{n}^\mu(\varepsilon t) \qquad \text{for all } \mu \in M,\ t \in [0, U/\varepsilon). \tag{2.44}$$

Proof. It is given in detail in Appendix A; however, here we sketch it in few lines. The main idea is to compare the inequalities (2.36) for $|\mathtt{L}(t)|^\mu$ and (2.43) for $\mathfrak{n}^\mu$, writing the second one with the change of variables $\tau = \varepsilon t$, $\tau' = \varepsilon t'$. The thesis follows using Lemma 2.5 with $\mathfrak{l}^\mu(t) := |\mathtt{L}(t)|^\mu$, $\mathfrak{v}^\mu(t) := \mathfrak{n}^\mu(\varepsilon t)$ and obvious choices for ξ^μ, η^μ; this is combined with a continuation principle for ODEs, to prove the existence of $(\mathtt{I}, \Theta)$ for all $t \in [0, U/\varepsilon)$. □

2G. A differential reformulation. We keep the assumptions at the beginning of the previous paragraph, but we require some more regularity on the functions $a^\mu, ..., e^\mu_{\nu\kappa}$, P^μ_ν, R^μ_ν fulfilling Eqs. (2.26)-(2.30) and (2.35), namely

$$a^\mu, b^\mu \in C^2(\Gamma_\rho, \mathbf{R}) \ , \qquad c^\mu, d^\mu_\nu, e^\mu_{\nu\kappa} \in C^1(\Gamma_\rho, \mathbf{R}) \ ,$$

$$R^\mu_\nu \in C^2([0,U), \mathbf{R}), \qquad P^\mu_\nu \in C^1([0,U), \mathbf{R}) \ . \tag{2.45}$$

2.7 ***Proposition.*** *(i) Assume there are* $\ell_* = (\ell^\mu_*) \in [0+\infty)^M$, $(A^\mu_\nu) \in [0, +\infty)^{M^2}$ *and* $\sigma = (\sigma^\mu) \in (0+\infty)^M$, *such that*

$$\Sigma := \Pi_{\mu \in M} [\ell^\mu_* - \sigma^\mu, \ell^\mu_* + \sigma^\mu] \subset \Pi_{\mu\in M}(0, \rho^\mu(0)/\varepsilon) \ , \tag{2.46}$$

$$\mathcal{A} := \max_{\mu \in M} \sum_{\nu \in M} A^\mu_\nu < 1/\varepsilon \ , \tag{2.47}$$

$$\left| \frac{\partial \alpha^\mu}{\partial r^\nu}(0, \varepsilon \ell) \right| \leqslant A^\mu_\nu \qquad \text{for } \mu, \nu \in M, \ \ell \in \Sigma, \tag{2.48}$$

$$|\alpha^\mu(0, \varepsilon\ell_*) - \ell^\mu_*| + \varepsilon A^\mu_\nu \sigma^\nu < \sigma^\mu \qquad \text{for } \mu \in M. \tag{2.49}$$

Then, there is a unique $\ell_0 = (\ell^\mu_0)$ *such that*

$$\ell_0 \in \Sigma \ , \qquad \alpha(0, \varepsilon\ell_0) = \ell_0 \ . \tag{2.50}$$

(ii) With ℓ_0 *as above, let* $\mathfrak{m} = (\mathfrak{m}^\mu), \mathfrak{n} = (\mathfrak{n}^\nu) \in C^1([0,U), \boldsymbol{R}^M)$ *solve the Cauchy problem*

$$\frac{d\mathfrak{m}^\mu}{d\tau} = P^\mu_\kappa \, \gamma^\kappa(\cdot, \varepsilon\mathfrak{n}, \mathfrak{n}) \ , \qquad \mathfrak{m}^\mu(0) = 0 \ , \tag{2.51}$$

$$\frac{d\mathfrak{n}^\mu}{d\tau} = \left(1 - \varepsilon \frac{\partial \alpha}{\partial r}(\cdot, \varepsilon\mathfrak{n})\right)^{-1,\mu}_{\ \ \lambda} \left(\frac{\partial \alpha^\lambda}{\partial \tau}(\cdot, \varepsilon\mathfrak{n}) + \varepsilon R^\lambda_\nu P^\nu_\kappa \, \gamma^\kappa(\cdot, \varepsilon\mathfrak{n}, \mathfrak{n}) + \varepsilon \frac{dR^\lambda_\nu}{d\tau} \mathfrak{m}^\nu \right) \ ,$$
$$\mathfrak{n}^\mu(0) = \ell^\mu_0 \tag{2.52}$$

($\mu \in M$) with the domain conditions

$$0 < \mathfrak{n}^\mu < \rho^\mu/\varepsilon \ , \qquad \det\left(1 - \varepsilon \frac{\partial \alpha}{\partial r}(\cdot, \varepsilon\mathfrak{n})\right) > 0 \tag{2.53}$$

(in the above, $1-\varepsilon\partial\alpha/\partial r$ stands for the matrix $(\delta^\mu_\nu - \varepsilon\partial\alpha^\mu/\partial r^\nu)$ $(\mu,\nu \in M)$, and Eq. (2.52) contains the matrix elements of its inverse. Note that (2.51) implies $\mathfrak{m}^\mu \geqslant 0$).
Then, the solution $(\mathtt{I},\Theta)$ of the perturbed system exists on $[0,U/\varepsilon)$ and

$$|\mathtt{L}(t)|^\mu \leqslant \mathfrak{n}^\mu(\varepsilon t) \qquad \text{for all } \mu \in M,\ t \in [0,U/\varepsilon). \tag{2.54}$$

Proof. See Appendix B, also containing a preliminary Lemma. □

2.8 Remark. The previous Proposition mentions ℓ_0, the unique fixed point of the map $\alpha(0,\varepsilon\cdot)$ in the set Σ of Eq. (2.46). The proof in the Appendix indicates that $\alpha(0,\varepsilon\cdot):\Sigma\to\Sigma$ has Lipschitz constant $\varepsilon\mathcal{A}<1$ in the maximum component norm $\|z\| := \max_\mu |z^\mu|$, with $\mathcal{A}$ as in (2.47). So, from the standard theory of contractions, we have the iterative construction

$$\ell_0 = \lim_{n\to+\infty} l_n, \qquad l_1 \text{ any point of } \Sigma\ ,\ l_n := \alpha(0,\varepsilon l_{n-1}) \qquad \text{for } n=2,3,\ldots\,. \tag{2.55}$$

For each $n \geqslant 2$,

$$\|\ell_0 - l_n\| \leqslant (\varepsilon\mathcal{A})^{n-1}\frac{\|l_2-l_1\|}{(1-\varepsilon\mathcal{A})}\ . \tag{2.56}$$

2H. Implementing the scheme in a typical case: the "$\mathcal{N}$-operation". Let us consider a situation in which (for given data (I_0,ϑ_0) and $\varepsilon>0$) we have analytical expressions for the solution $\mathtt{J}$ of the averaged system (1.8) (on an interval $[0,U)$) and for all the auxiliary functions $\mathtt{R},\mathtt{K},s,p,\ldots,\mathscr{G},\mathscr{H}$, $a^\mu,\ldots,e^\mu_{\nu\kappa}$, P^μ_ν, R^μ_ν of the previous paragraphs (having chosen a separating family of seminorms $(|\ |^\mu)_{\mu\in M}$, and making the regularity assumptions (2.45)). One can provide nontrivial examples where these expressions can be obtained: one of them is considered in Section 3.

In this situation, to obtain the final estimates $|\mathtt{L}(t)|^\mu \leqslant \mathfrak{n}^\mu(\varepsilon t)$ of Proposition 2.7 we need: the fixed point ℓ_0, defined by Eq. (2.50); the functions $\mathfrak{m}=(\mathfrak{m}^\mu),\mathfrak{n}=(\mathfrak{n}^\mu)$ fulfilling the Cauchy problem (2.51) (2.52). Typically, to find ℓ_0 and $\mathfrak{m}^\mu,\mathfrak{n}^\nu$ analytically will be difficult or impossible, and a numerical approach will be required. Concerning ℓ_0, one can compute numerically the iterates $l_2,l_3,\ldots l_n$ in (2.55) up to a sufficiently large order n, and then approximate ℓ_0 with l_n. As for $\mathfrak{m}^\mu,\mathfrak{n}^\nu$, one can attack the Cauchy problem (2.51) (2.52) by any package for the numerical integration of ODEs (paying attention to the domain conditions (2.53)).

From now on, the term "$\mathcal{N}$-operation" will be employed to indicate the numerical determination of $\ell_0,\mathfrak{m},\mathfrak{n}$ along the above lines ([3]). Generally, the $\mathcal{N}$-operation to find ℓ_0 and $\mathfrak{m},\mathfrak{n}$ on the interval $[0,U)$ is faster (and more reliable) than the computation of $\mathtt{L}(t) := [\mathtt{I}(t)-\mathtt{J}(\varepsilon t)]/\varepsilon$ on the long interval $t\in[0,U/\varepsilon)$, through a direct numerical

[3] this is somehow different from the "$\mathfrak{N}$-operation" of [1], that also included the numerical computation of $\mathtt{J},\mathtt{R},\mathtt{K}$ on $[0,U)$ and was designed to work with a single norm on $\mathbf{R}^d$.

attack to the perturbed system (1.5): we think that this gives a practical value to the general framework developed here. This situation will be exemplified in Section 3.

2I. The "$\mathcal{L}$-operation". This expression will be used to indicate the direct numerical computation of $\mathtt{L}$ from the perturbed system (1.5), on the time interval $[0, U/\varepsilon)$; in the general framework of this paper, this operation must be performed only if one wants to test the efficiency of the $\mathcal{N}$-operation.

Let us clarify the previous statements, assuming again to have the analytical expressions of all the functions mentioned in paragraph 2H. Having the expression of $\mathtt{J}$, we substitute $\mathtt{I}(t) = \mathtt{J}(\varepsilon t) + \varepsilon \mathtt{L}(t)$ in Eqs. (1.5) for $(\mathtt{I}, \Theta)$; this gives rise to the Cauchy problem

$$\begin{cases} (d\mathtt{L}/dt)(t) = f(\mathtt{J}(\varepsilon t) + \varepsilon \mathtt{L}(t), \Theta(t)) - \overline{f}(\mathtt{J}(\varepsilon t)), & \mathtt{L}(0) = 0 \ , \\ (d\Theta/dt)(t) = \omega(\mathtt{J}(\varepsilon t)) + \varepsilon g(\mathtt{J}(\varepsilon t) + \varepsilon \mathtt{L}(t), \Theta(t)), & \Theta(0) = \vartheta_0 \end{cases} \tag{2.57}$$

for the unknown functions $t \mapsto (\mathtt{L}(t), \Theta(t))$. By definition, the "$\mathcal{L}$-operation" is the numerical solution of (2.57) for $t \in [0, U/\varepsilon)$ ([4]).

The efficiency of the $\mathcal{N}$ operation is tested via $\mathcal{L}$ comparing: (1) the CPU times $\mathfrak{T}_{\mathcal{L}}$, $\mathfrak{T}_{\mathcal{N}}$ required to perform both operations on standard machines; (2) the graphs of the functions $|\mathtt{L}^\mu|$ and of their estimators $\mathfrak{n}^\mu$, made available by the two operations. Of course, the test is satisfactory if:

(i) $\mathfrak{T}_{\mathcal{N}}$ is considerably shorter than $\mathfrak{T}_{\mathcal{L}}$;

(ii) for each $\mu \in M$ the estimator $t \mapsto \mathfrak{n}^\mu(\varepsilon t)$ approximates well the envelope of the rapidly oscillating function $t \mapsto |\mathtt{L}(t)|^\mu$, for $t \in [0, U/\varepsilon)$.

The whole procedures concerning $\mathcal{N}$ and $\mathcal{L}$ are illustrated in the next section; in the example therein, both (i) and (ii) will occur in the test of $\mathcal{N}$ via $\mathcal{L}$.

[4] This differs from the "$\mathfrak{L}$-operation" of [1], which included a preliminary numerical determination of $\mathtt{J}$ on $[0, U)$.

3 An example from rigid body dynamics.

3A. Introducing the example. We consider a perturbed integrable system of the form (1.5), with

$$d = 2, \qquad \Lambda := \{I = (I^1, I^2) \mid I^1, I^2 \in (0, +\infty)\}, \qquad \omega(I) = I^1 I^2, \tag{3.1}$$

$$f(I, \vartheta) := \big(- I^1(\lambda_1 + \mu \cos(2\vartheta)), -I^2(\lambda_2 - \mu \cos(2\vartheta)\big) \ , \quad g(I, \vartheta) := \mu \sin(2\vartheta) \ ;$$

this depends on three real coefficients $\mu, \lambda_1, \lambda_2$ such that

$$\lambda_1 > 0 \ , \qquad -\lambda_1 < \mu < \lambda_1 \ , \qquad \lambda_2 > -\lambda_1 \ . \tag{3.2}$$

This system has already appeared in [1] (Section 4, Example 4) where it was related to Euler's equation for a rigid body with gyroscopic symmetry under a damping moment proportional to ε, with a particular dependence on the angular velocity. The actions I^1, I^2 have different physical meaning: in fact, I^1 is the equatorial angular velocity (in suitable units) and I^2 measures the inclination of the angular velocity on the gyroscopic axis ([5]). We will take for (1.5) the initial conditions

$$I_0 = (I_0^1, I_0^2) \in \Lambda \ , \qquad \vartheta_0 := 0 \ . \tag{3.3}$$

The forthcoming analysis shows that, depending on the data and on the other parameters involved in the problem, the numerical values of the solution components $\mathtt{I}^i$ $(i = 1, 2)$ can be very different over long times; the same happens for the components $\mathtt{J}^i$ and $\mathtt{L}^i(t) = [\mathtt{I}^i(t) - \mathtt{J}^i(\varepsilon t)]/\varepsilon$. For this reasons, and for the different meaning of the two actions, it can be of interest to derive separate estimates for the absolute values $|\mathtt{L}^i(t)|$ $(i = 1, 2)$; this will mark a difference with the analysis of [1], where we only gave a global estimate for $\sqrt{(L^1)^2 + (\mathtt{L}^2)^2}$.

3B. Analysis of the example. The average $\overline{f}$ and the solutions $\mathtt{J}$ of (1.8), $\mathtt{R}, \mathtt{K}$ of (2.11) (2.12) (on any interval $[0, U)$) are written in the forthcoming Table 1, which also reports the auxiliary functions $s, ..., \mathscr{H}$ required by our method. As anticipated, our aim is to estimate separately the absolute values $|\mathtt{L}^i|$ $(i = 1, 2)$; this marks the difference with respect to [1], where this example was treated estimating the Euclidean norm $\sqrt{(\mathtt{L}^1)^2 + (\mathtt{L}^2)^2}$. In the language of Section 1, analyzing the components of $\mathtt{L}$ corresponds to use the seminorms (1.16), i.e.,

$$| \ |^i : \mathbf{R}^2 \to [0, +\infty) \ , \qquad X \mapsto |X|^i := |X^i| \qquad\qquad (i = 1, 2) \ . \tag{3.4}$$

Whenever necessary, we will use for $T^1_1(\mathbf{R}^2)$, $T^1_2(\mathbf{R}^2)$ the consistent seminorms (2.8).

The second half of Table 1 contains the functions $\rho^i, a^i, ..., d^i_j, e^i_{jk}, R^i_j, P^i_j$ $(i, j, k = 1, 2)$ required by the general framework of the previous section. The choice of ρ^i is an

[5] See Eq. (4.18) of the cited work.

obvious consequence of the form of Λ; a^i, b^i, c^i have been computed binding the left-hand sides of Eqs. (2.26)-(2.28) by elementary means (similar to the ones employed for the Euclidean norm estimates of [1], but here applied to the components); we observe that $a^i(\tau, r)$ is just a bound on $|s(\mathtt{J}(\tau), \delta J)|^i$ (for $|\delta J|^k = r^k$), because $\mathtt{K}(\tau)$ and $s(I_0, \vartheta_0)$ in the left-hand side of (2.26) are zero. The functions d^i_j, e^i_{jk} are identically zero due to the vanishing of $\mathscr{G}, \mathscr{H}$; the expressions for R^i_j and P^i_j are just the ones of the matrix elements of $\mathtt{R}$ and $\mathtt{R}^{-1}$.

3C. Results. Starting from the functions in Table 1, the $\mathcal{N}$-operation has been performed for two choices of the initial data I^1_0, I^2_0 and of the parameters U, ε, $\mathfrak{l}_1, \mathfrak{l}_2, \mu$, producing as a main output the estimators $\mathfrak{n} = (\mathfrak{n}^i)_{i=1,2}$ on $[0, U)$; the $\mathcal{L}$-operation has been performed as a test, producing directly the function $\mathtt{L} = (L^i)_{i=1,2}$. Both operations were carried over on a PC, using the MATHEMATICA package.

The results are summarized in Figures a,b,c,d. These report the CPU times $\mathfrak{T}_{\mathcal{N}}$, $\mathfrak{T}_{\mathcal{L}}$ (in seconds) and allow to compare the actual values of $|\mathtt{L}^i(t)|$ with our estimators $\mathfrak{n}^i(\tau)$, for $t = \tau/\varepsilon$ and $\tau \in [0, U)$.

It turns out that $\mathfrak{T}_{\mathcal{N}}/\mathfrak{T}_{\mathcal{L}} \simeq 1/12$ for the first choice of the parameters (Figures a, b) and $\mathfrak{T}_{\mathcal{N}}/\mathfrak{T}_{\mathcal{L}} \simeq 1/600$ for the second choice (Figures c, d), where the time scale U/ε is overwhelmingly long. In all cases the functions $\tau \mapsto \mathfrak{n}^i(\tau)$ practically coincide with the envelopes of the oscillating functions $\tau \mapsto |\mathtt{L}^i(\tau/\varepsilon)|$, for $\tau \in [0, U)$.

Table 1. A list of functions for the example.

For $I = (I^1, I^2) \in (0, +\infty)^2$, $\vartheta \in \mathbf{T}$ and $\delta I = (\delta I^1, \delta I^2) \in (-I^1, +\infty) \times (-I^2, +\infty)$:

$$\overline{f}(I) = (-\lambda_1 I^1, -\lambda_2 I^2) \ ;$$

$$s(I,\vartheta) = \frac{\mu}{2}\sin(2\vartheta)\left(-\frac{1}{I^2}\ , \frac{1}{I^1}\right), \qquad v(I,\vartheta) = \frac{\mu}{2I^1 I^2}\sin^2\vartheta\left(-\frac{1}{I^2}, \frac{1}{I^1}\right) \ ,$$

$$p(I,\vartheta) = \frac{\mu\sin(2\vartheta)}{2}\left(-\frac{\lambda_2 + \mu\cos(2\vartheta)}{I^2}, \frac{\lambda_1 + 3\mu\cos(2\vartheta)}{I^1}\right), \qquad \overline{p}(I) = (0,0) \ ,$$

$$q(I,\vartheta) = \frac{\mu\sin^2\vartheta}{2I^1 I^2}\left(-\frac{2\lambda_2 + 2\mu + \lambda_1 + \mu\cos(2\vartheta)}{I^2}, \frac{\lambda_2 + 2\mu + 2\lambda_1 + 3\mu\cos(2\vartheta)}{I^1}\right),$$

$$w(I,\vartheta) = \frac{\mu\sin^2\vartheta}{2I^1 I^2}\left(-\frac{\lambda_2 + \mu\cos^2\vartheta}{I^2}, \frac{\lambda_1 + 3\mu\cos^2\vartheta}{I^1}\right) \ , \qquad u(I,\vartheta) = \frac{\mu\sin^2\vartheta}{4\, I^1 I^2}$$

$$\times\Bigg(- \frac{4\lambda_2^2 + 6\lambda_2\mu + 2\lambda_2\lambda_1 + \mu\lambda_1 + \mu(4\lambda_2 + 3\mu + \lambda_1)\cos(2\vartheta) + 3\mu^2\cos^2(2\vartheta)}{I^2} \ ,$$

$$\frac{3\lambda_2\mu + 2\lambda_2\lambda_1 + 10\mu\lambda_1 + 4\lambda_1^2 + 3\mu(\lambda_2 + 5\mu + 4\lambda_1)\cos(2\vartheta) + 15\mu^2\cos^2(2\vartheta)}{I^1}\Bigg) \ ;$$

$$\frac{\partial\overline{f}}{\partial I}(I) = \mathrm{diag}(-\lambda_1, -\lambda_2) \ , \ \mathscr{M}(I) = \mathrm{diag}(-\lambda_1^2, -\lambda_2^2) \ , \ \mathscr{G}(I, \delta I) = 0, \ \mathscr{H}(I, \delta I) = 0 \ .$$

For $\tau \in [0, U)$:

$$\mathtt{J}(\tau) = (I_0^1\, e^{-\lambda_1\tau}, I_0^2\, e^{-\lambda_2\tau}) \ ; \quad \mathtt{R}(\tau) = \mathrm{diag}(e^{-\lambda_1\tau}, e^{-\lambda_2\tau}) \ , \quad \mathtt{K}(\tau) = (0,0) \ ;$$

$$\rho^i(\tau) := \mathtt{J}^i(\tau) \ (i = 1,2).$$

For $\tau \in [0, U)$ and $r = (r^1, r^2) \in [0, \mathtt{J}^1(\tau)) \times [0, \mathtt{J}^2(\tau))$:

$$a^1(\tau, r) := \frac{|\mu|}{2(\mathtt{J}^2(\tau) - r^2)} \ , \qquad a^2(\tau, r) := \frac{|\mu|}{2(\mathtt{J}^1(\tau) - r^1)} \ ;$$

$$b^1(\tau, r) := \ \frac{|\mu|(4\lambda_1 + 4\lambda_2 + |\mu|)}{8(\mathtt{J}^1(\tau) - r^1)(\mathtt{J}^2(\tau) - r^2)^2} \ , \qquad b^2(\tau, r) := \ \frac{|\mu|(4\lambda_1 + 4\lambda_2 + 3|\mu|)}{8(\mathtt{J}^1(\tau) - r^1)^2(\mathtt{J}^2(\tau) - r^2)} \ ;$$

$$c^1(\tau, r) := \ \frac{|\mu|(16(\lambda_1 + \lambda_2)^2 + 16|\mu|(\lambda_1 + \lambda_2) + 3|\mu|^2)}{16(\mathtt{J}^1(\tau) - r^1)(\mathtt{J}^2(\tau) - r^2)^2} \ ,$$

$$c^2(\tau, r) := \ \frac{|\mu|(16(\lambda_1 + \lambda_2)^2 + 16|\mu|(\lambda_1 + \lambda_2) + 15|\mu|^2)}{16(\mathtt{J}^1(\tau) - r^1)^2(\mathtt{J}^2(\tau) - r^2)} \ ;$$

$$d^i_j(\tau, r) := 0 \ ; \qquad e^i_{jk}(\tau, r) := 0 \ ;$$

$$R^i_j(\tau) := e^{-\lambda_i\tau}\delta^i_j \ ; \qquad P^i_j(\tau) := e^{\lambda_i\tau}\delta^i_j \qquad\qquad (i, j, k = 1, 2) \ .$$

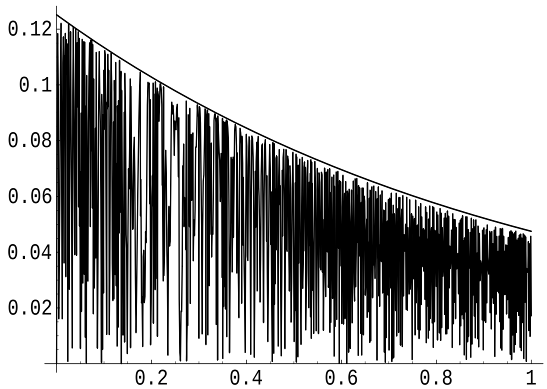

Figure a. $\mu=1$, $\lambda_1=2$, $\lambda_2=-1$, $I_0^1=4$, $I_0^2=4$, $\varepsilon=10^{-2}$, $U=1$. $\mathfrak{T}_{\mathcal{N}}=0.031s$, $\mathfrak{T}_{\mathfrak{L}}=0.391s$. Graphs of $\mathfrak{n}^1(\tau)$ and $|\mathtt{L}^1(\tau/\varepsilon)|$.

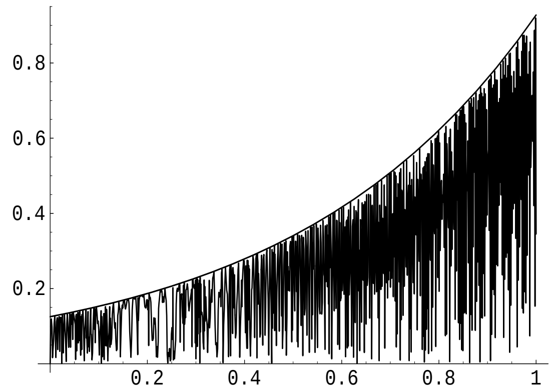

Figure b. The same parameters as in Fig.a. Graphs of $\mathfrak{n}^2(\tau)$ and $|\mathtt{L}^2(\tau/\varepsilon)|$.

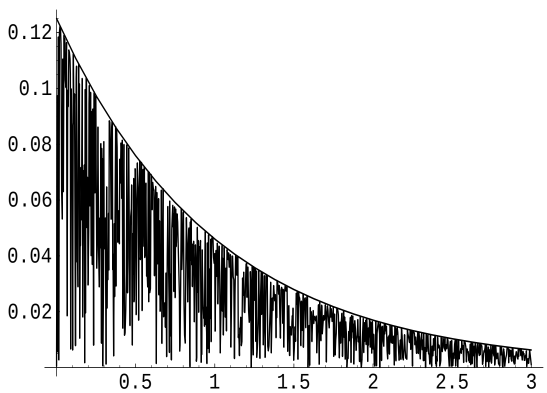

Figure c. $\mu=1$, $\lambda_1=1.1$, $\lambda_2=-1$, $I_0^1=4$, $I_0^2=4$, $\varepsilon=10^{-3}$, $U=3$. $\mathfrak{T}_{\mathcal{N}}=0.047s$, $\mathfrak{T}_{\mathfrak{L}}=31.3s$. Graphs of $\mathfrak{n}^1(\tau)$ and $|\mathtt{L}^1(\tau/\varepsilon)|$.

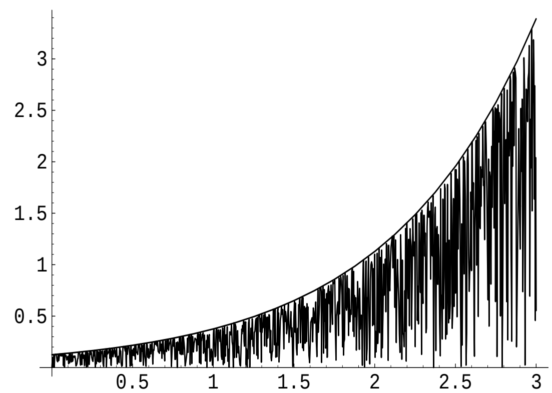

Figure d. The same parameters as in Fig.c. Graphs of $\mathfrak{n}^2(\tau)$ and $|\mathtt{L}^2(\tau/\varepsilon)|$.

A Appendix. Proof of Proposition 2.6.

Let $[0, V/\varepsilon)$ (with $V \in [0, +\infty]$) be the domain of the maximal solution $(\mathtt{I}, \Theta)$ of (1.5), and put

$$U' := \min(V, U) \ . \tag{A.1}$$

Step 1. One has $|\mathtt{L}(t)|^\mu < \mathfrak{n}^\mu(\varepsilon t)$ for all $\mu \in M$, $t \in [0, U'/\varepsilon)$. To show this, we write the integral inequality (2.43) with $\tau = \varepsilon t$, $\tau' = \varepsilon t'$; this gives

$$\mathfrak{n}^\mu(\varepsilon t) > \alpha^\mu(\varepsilon t, \varepsilon \mathfrak{n}(\varepsilon t)) + \varepsilon^2 R^\mu_\lambda(\varepsilon t) \int_0^{\varepsilon t} dt' P^\lambda_\nu(\varepsilon t')\, \gamma^\nu(\varepsilon t', \varepsilon \mathfrak{n}(\varepsilon t'), \mathfrak{n}(\varepsilon t')) \tag{A.2}$$

for $t \in [0, U/\varepsilon)$, hence for $t \in [0, U'/\varepsilon)$.

To go on, we use Lemma 2.4 with the constant U therein replaced by U'; this gives Eq. (2.36) for $t \in [0, U'/\varepsilon)$. Keeping in mind Eqs. (2.36) and (A.2), we apply Lemma 2.5 with $T := U'/\varepsilon$, $\delta^\mu(t) := \rho^\mu(\varepsilon t)/\varepsilon$, $\xi^\mu(t, \ell) := \alpha^\mu(\varepsilon t, \varepsilon \ell)$, $\eta^\mu(t, t', \ell) := \varepsilon^2 R^\mu_\nu(\varepsilon t) P^\nu_\kappa(\varepsilon t') \gamma^\kappa(\varepsilon t', \varepsilon \ell, \ell)$, $\mathfrak{l}^\mu(t) := |\mathtt{L}(t)|^\mu$, $\mathfrak{v}^\mu(t) := \mathfrak{n}^\mu(\varepsilon t)$ (the requirement $\mathfrak{l}^\mu(0) = 0$ is fulfilled by construction). Lemma 2.5 gives $\mathfrak{l}^\mu(t) < \mathfrak{v}^\mu(t)$, yielding the thesis.

Step 2. One has $V \geqslant U$, i.e., $U' = U$ (thus $(\mathtt{I}, \Theta)$ exists on $[0, U/\varepsilon)$ and the inequality of Step 1 holds in this interval). Indeed, suppose $V < U$ and put

$$K := \{(t, I) \in [0, V/\varepsilon] \times \mathbf{R}^d \mid |I - \mathtt{J}(\varepsilon t)|^\mu \leqslant \varepsilon \mathfrak{n}^\mu(\varepsilon t) \quad \forall \mu \in M\} \ . \tag{A.3}$$

This is a compact subset of $\mathbf{R} \times \mathbf{R}^d$ and (due to Step 1 and (2.24)) graph $(\mathtt{I}, \Theta) \subset K \times \mathbf{T} \subset \mathbf{R} \times \Lambda \times \mathbf{T}$. But $K \times \mathbf{T}$ is compact: so, by the continuation principle for ODEs [7], the solution $(\mathtt{I}, \Theta)$ can be extended to an interval larger than $[0, V/\varepsilon)$. This contradicts our maximality assumption, and concludes the proof. $\square$

B Appendix. Proof of Proposition 2.7.

We begin with a Lemma, holding under the assumptions at the beginning of paragraph 2E. Its proof is very similar to the one given in [1] for Lemma C.1, so it will not be reported.

B.1 ***Lemma.*** *Assume there are functions $\mathfrak{n}_\delta = (\mathfrak{n}^\mu_\delta) \in C([0, U_\delta), [0, +\infty)^M)$, labeled by a parameter $\delta \in (0, \delta_*]$, such that the following holds:*
(i) $U_\delta \to U$ for $\delta \to 0^+$;
(ii) for all $\delta \in (0, \delta_*]$, $\mu \in M$ and $\tau \in [0, U_\delta)$,

$$\mathfrak{n}^\mu_\delta(\tau) < \rho^\mu(\tau)/\varepsilon \ , \tag{B.1}$$

$$\mathfrak{n}^\mu_\delta(\tau) = \delta + \alpha^\mu(\tau, \varepsilon\mathfrak{n}_\delta(\tau)) + \varepsilon R^\mu_\lambda(\tau)\int_0^\tau d\tau' P^\lambda_\kappa(\tau')\gamma^\kappa(\tau', \varepsilon\mathfrak{n}_\delta(\tau'), \mathfrak{n}_\delta(\tau')) \ ; \qquad \text{(B.2)}$$

(iii) for each fixed $\tau \in [0, U)$, the limit

$$\mathfrak{n}(\tau) := \lim_{\delta\to 0^+} \mathfrak{n}_\delta(\tau) \qquad \text{(B.3)}$$

exists in $[0, +\infty)^M$.

Then the solution $(\mathtt{I}, \Theta)$ of (1.5) exists in $[0, U/\varepsilon)$ and, defining $\mathtt{L}$ as in Eq. (2.13),

$$|\mathtt{L}(t)|^\mu \leqslant \mathfrak{n}^\mu(\varepsilon t) \qquad \textit{for all } \mu \in M,\ t \in [0, U/\varepsilon). \qquad \text{(B.4)}$$

We come to Proposition 2.7; so, we make the assumptions at the beginning of paragraph 2G, strengthened by the smoothness requirements (2.45) for the functions $a^\mu, ..., e^\mu_{\nu\kappa}$.

Proof of Proposition 2.7. We introduce the norm $\| \ \|$ on $\mathbf{R}^M$, the function α_0 and its generalizations α_δ ($\delta \in [0, +\infty)$) setting

$$\|z\| := \max_{\mu\in M} |z^\mu| \ ; \qquad \text{(B.5)}$$

$$\alpha_0 : \Sigma \to \mathbf{R}, \qquad \ell \mapsto \alpha_0(\ell) := \alpha(0, \varepsilon\ell) \ , \qquad \text{(B.6)}$$

$$\alpha_\delta : \Sigma \to \mathbf{R} \ , \qquad \ell \mapsto \alpha_\delta(\ell) := \alpha_0(\ell) + \delta \ . \qquad \text{(B.7)}$$

Step 1. For each $\delta \geqslant 0$, α_δ is a contractive map with respect to the norm $\| \ \|$. In fact, by (2.48) it is

$$\left|\frac{\partial\alpha^\mu_\delta}{\partial\ell^\nu}(\ell)\right| = \varepsilon\left|\frac{\partial\alpha^\mu}{\partial r^\nu}(0, \varepsilon\ell)\right| \leqslant \varepsilon A^\mu_\nu \qquad \text{for all } \ell \in \Sigma \ . \qquad \text{(B.8)}$$

Let $\ell, \ell' \in \Sigma$; the equation $\alpha^\mu_\delta(\ell) - \alpha^\mu_\delta(\ell') = \int_0^1 ds (\partial\alpha^\mu_\delta/\partial r^\nu)((1-s)\ell' + s\ell) \times (\ell^\nu - \ell'^\nu)$ implies

$$|\alpha^\mu_\delta(\ell) - \alpha^\mu_\delta(\ell')| \leqslant \varepsilon A^\mu_\nu |\ell^\nu - \ell'^\nu| \ , \qquad \text{(B.9)}$$

whence

$$||\alpha_\delta(\ell) - \alpha_\delta(\ell')|| \leqslant \varepsilon\mathcal{A}||\ell - \ell'|| < ||\ell - \ell'|| \ . \qquad \text{(B.10)}$$

The last two inequalities depend on (2.47); contractivity of α_δ is proved.
Step 2. There is $\delta_* > 0$ such that, for all $\delta \in [0, \delta_*]$, α_δ sends Σ into itself. In fact, for any $\delta \geqslant 0$ and $\ell \in \Sigma$,

$$|\alpha^\mu_\delta(\ell) - \ell^\mu_*| = |\alpha^\mu_0(\ell) + \delta - \ell^\mu_*| \leqslant |\alpha^\mu_0(\ell) - \alpha^\mu_0(\ell_*)| + |\alpha^\mu_0(\ell_*) - \ell^\mu_*| + \delta$$

$$\leqslant \varepsilon A^\mu_\nu |\ell^\nu - \ell^\nu_*| + |\alpha^\mu_0(\ell_*) - \ell^\mu_*| + \delta \leqslant \varepsilon A^\mu_\nu \sigma^\nu + |\alpha^\mu_0(\ell_*) - \ell^\mu_*| + \delta \ , \qquad \text{(B.11)}$$

where the second inequality follows from Eq. (B.9) with $\delta = 0$. To go on, we note that (2.49) implies the existence of a $\delta_* > 0$ such that

$$|\alpha^\mu(0, \varepsilon \ell_*) - \ell^\mu_*| + \varepsilon A^\mu_\nu \sigma^\nu + \delta_* \leqslant \sigma^\mu \qquad \text{for all } \mu \in M. \tag{B.12}$$

For $\delta \in [0, \delta_*]$ and $\ell \in \Sigma$, Eqs. (B.11), (B.12) imply $|\alpha^\mu_\delta(\ell) - \ell^\mu_*| \leqslant \sigma^\mu$, i.e., $\alpha_\delta(\ell) \in \Sigma$.
Step 3. For all $\delta \in [0, \delta_*]$, the map α_δ has a unique fixed point $\ell_\delta = (\ell^\mu_\delta) \in \Sigma$, which depends continuously on δ. Existence and uniqueness of the fixed point follows from the Banach theorem on contractions; to prove continuity we note that, for all $\delta, \delta' \in [0, \delta_*]$,

$$||\ell_\delta - \ell_{\delta'}|| = ||\alpha_\delta(\ell_\delta) - \alpha_{\delta'}(\ell_{\delta'})|| = ||\alpha_0(\ell_\delta) + \delta - \alpha_0(\ell_{\delta'}) - \delta'|| \tag{B.13}$$

$$\leqslant ||\alpha_0(\ell_\delta) - \alpha_0(\ell_{\delta'})|| + |\delta - \delta'| \leqslant \varepsilon \mathcal{A} ||\ell_\delta - \ell_{\delta'}|| + |\delta - \delta'| \ ,$$

the last inequality depending on (B.10) with $\delta = 0$. This implies

$$||\ell_\delta - \ell_{\delta'}|| \leqslant \frac{|\delta - \delta'|}{1 - \varepsilon \mathcal{A}} \ ; \tag{B.14}$$

so the map $\delta \mapsto \ell_\delta$ is Lipschitz, and a fortiori continuous.
Step 4. Proving the thesis of (i). This follows from Step 3, with $\delta = 0$.
Step 5. Proving the thesis of (ii). For any $\delta \in [0, \delta_*]$, let ℓ_δ be as in Step 3. First of all, let us consider the space of real matrices $N = (N^\mu_\nu)_{\mu,\nu \in M}$, with the norm $\|N\| := \sup_{z \in \mathbf{R}^M, z \neq 0} \|Nz\| / \|z\|$, and note that the second inequality (B.8), with (2.47), yields

$$\varepsilon \| \frac{\partial \alpha}{\partial r}(0, \ell_\delta) \| \leqslant \varepsilon \mathcal{A} < 1 \ . \tag{B.15}$$

By a well-known fact on matrices of the form $1 - N$, this implies

$$\det(1 - \varepsilon \frac{\partial \alpha}{\partial r}(0, \ell_\delta)) > 0 \ . \tag{B.16}$$

From the standard continuity theorems for the solutions of a parameter-dependent Cauchy problem, we know that there is a family $(U_\delta, \mathfrak{m}_\delta, \mathfrak{n}_\delta)_{\delta \in (0, \delta_*]}$ with the forthcoming properties (a) (b):
(a) for all $\delta \in (0, \delta_*]$, it is $\mathfrak{m}_\delta = (\mathfrak{m}^\mu_\delta), \mathfrak{n}_\delta = (\mathfrak{n}^\mu_\delta) \in C^1([0, U_\delta), \mathbf{R}^M)$; furthermore,

$$\frac{d \mathfrak{m}^\mu_\delta}{d\tau} = P^\mu_\kappa \gamma^\kappa(\cdot, \varepsilon \mathfrak{n}_\delta, \mathfrak{n}_\delta) \ , \qquad \mathfrak{m}^\mu_\delta(0) = 0 \ , \tag{B.17}$$

$$\frac{d \mathfrak{n}^\mu_\delta}{d\tau} = \Big(1 - \varepsilon \frac{\partial \alpha}{\partial r}(\cdot, \varepsilon \mathfrak{n}_\delta)\Big)^{-1,\mu}_{\lambda} \left(\frac{\partial \alpha^\lambda}{\partial \tau}(\cdot, \varepsilon \mathfrak{n}_\delta) + \varepsilon R^\lambda_\nu P^\nu_\kappa \, \gamma^\kappa(\cdot, \varepsilon \mathfrak{n}_\delta, \mathfrak{n}_\delta) + \varepsilon \frac{d R^\lambda_\nu}{d\tau} \mathfrak{m}^\nu_\delta \right),$$
$$\mathfrak{n}^\mu_\delta(0) = \ell^\mu_\delta \tag{B.18}$$

$$0 < \mathfrak{n}^\mu_\delta < \rho^\mu/\varepsilon\ , \qquad \det(1 - \varepsilon\frac{\partial\alpha}{\partial r}(\cdot, \varepsilon\mathfrak{n}_\delta)) > 0 \tag{B.19}$$

(concerning the domain conditions in the last line, recall (B.16); $\mathfrak{m}^\mu_\delta \geqslant 0$ by (B.17)).
(b) One has

$$U_\delta \to_{\delta\to 0^+} U, \quad \mathfrak{n}^\mu_\delta(t) \to_{\delta\to 0^+} \mathfrak{n}^\mu(t), \quad \mathfrak{m}^\mu_\delta(t) \to_{\delta\to 0^+} \mathfrak{m}^\mu(t) \ \text{ for all } t \in [0, U), \tag{B.20}$$

where U, $\mathfrak{m}^\mu$, $\mathfrak{n}^\mu$ are as stated in (ii).

Let us consider the pair $\mathfrak{m}_\delta, \mathfrak{n}_\delta$ for any $\delta \in (0, \delta_*]$. Then, integrating (B.17),

$$\mathfrak{m}^\mu_\delta(\tau) = \int_0^\tau d\tau' P^\mu_\nu(\tau')\gamma^\nu(\tau', \varepsilon\mathfrak{n}_\delta(\tau'), \mathfrak{n}_\delta(\tau')) \qquad \text{for } \tau \in [0, U_\delta). \tag{B.21}$$

Furthermore, from Eq. (B.18) and (B.17) we infer

$$\begin{aligned}
0 &= \Big(1 - \varepsilon\frac{\partial\alpha}{\partial r}(\cdot, \varepsilon\mathfrak{n}_\delta)\Big)^\mu_{\ \varsigma}\frac{d\mathfrak{n}^\varsigma_\delta}{d\tau} - \left(\frac{\partial\alpha^\mu}{\partial\tau}(\cdot, \varepsilon\mathfrak{n}_\delta) + \varepsilon R^\mu_\nu P^\nu_\kappa\,\gamma^\kappa(\cdot, \varepsilon\mathfrak{n}_\delta, \mathfrak{n}_\delta) + \varepsilon\frac{dR^\mu_\nu}{d\tau}\,\mathfrak{m}^\nu_\delta\right) \\
&= \frac{d\mathfrak{n}^\mu_\delta}{d\tau} - \varepsilon\frac{\partial\alpha^\mu}{\partial r^\varsigma}(\cdot, \varepsilon\mathfrak{n}_\delta)\frac{d\mathfrak{n}^\varsigma_\delta}{d\tau} - \left(\frac{\partial\alpha^\mu}{\partial\tau}(\cdot, \varepsilon\mathfrak{n}_\delta) + \varepsilon R^\mu_\nu\frac{d\mathfrak{m}^\nu_\delta}{d\tau} + \varepsilon\frac{dR^\mu_\nu}{d\tau}\,\mathfrak{m}^\nu_\delta\right) \\
&= \frac{d}{d\tau}(\mathfrak{n}^\mu_\delta - \alpha^\mu(\cdot, \varepsilon\mathfrak{n}_\delta) - \varepsilon R^\mu_\nu\mathfrak{m}^\nu_\delta)\ ;
\end{aligned} \tag{B.22}$$

therefore, for $\tau \in [0, U)$,

$$\begin{aligned}
\mathfrak{n}^\mu_\delta(\tau) - \alpha^\mu(\tau, \varepsilon\mathfrak{n}_\delta(\tau)) - \varepsilon R^\mu_\nu(\tau)\,\mathfrak{m}^\nu_\delta(\tau) &= \mathfrak{n}^\mu_\delta(0) - \alpha^\mu(0, \varepsilon\mathfrak{n}_\delta(0)) - \varepsilon R^\mu_\nu(0)\,\mathfrak{m}^\nu_\delta(0) \\
&= \ell^\mu_\delta - \alpha^\mu(0, \varepsilon\ell_\delta) = \ell^\mu_\delta - \alpha^\mu_0(\ell_\delta) = \delta
\end{aligned} \tag{B.23}$$

(recall the initial conditions in Eqs. (B.17) (B.18), Eqs. (B.6) (B.7) and Step 3, giving $\ell^\mu_\delta = \alpha^\mu_\delta(\ell_\delta) = \alpha^\mu_0(\ell_\delta) + \delta$).

From Eqs. (B.23) (B.21) we see that $\mathfrak{n}_\delta$ fulfils Eq. (B.2) of Lemma B.1. Due to Eq. (B.20) on the limit for $\delta \to 0^+$, from the cited Lemma we obtain the thesis. □

Acknowledgments. This work has been partially supported by the GNFM of Istituto Nazionale di Alta Matematica and by MIUR, Research Project Cofin/2006 "Metodi geometrici nella teoria delle onde non lineari e applicazioni".

References

[1] C. Morosi, L. Pizzocchero, *On the average principle for one-frequency systems*, J. Phys. A: Math. Gen. **39**, 3673-3702 (2006).

[2] C. Morosi, L. Pizzocchero, *On the averaging principle for one-frequency systems. An application to satellite motions*, arXiv:math-ph/0609038.

[3] V.I. Arnold, *Mathematical methods of classical mechanics*, Springer-Verlag, New York (1978); J.A. Sanders, F. Verhulst, *Averaging methods in nonlinear dynamical systems*, Springer-Verlag, New York (1985); P. Lochak, C. Meunier, *Multiphase averaging for classical systems*, Springer-Verlag, New York (1988).

[4] D.R. Smith, *Singular perturbation theory*, Cambridge Univ. Press, Cambridge (1985).

[5] J. Horwath, *Topological vector spaces and distributions. I*, Addison-Wesley, Reading (Massachusetts) (1966).

[6] D.S. Mitrinovic, J.E. Pecaric, A.M. Fink, *Inequalities involving functions and their integrals and derivatives*, Kluwer, Dordrecht (1991).

[7] E. Zeidler, *Nonlinear functional analysis and its applications II/B. Nonlinear monotone operators*, Springer-Verlag, New York (1990).